

Quantum Hamiltonian identification from measurement time traces

Jun Zhang¹ and Mohan Sarovar^{2*}

¹*Joint Institute of UMich-SJTU and Key Laboratory of System Control and Information Processing (MOE), Shanghai Jiao Tong University, Shanghai, 200240, China*

²*Department of Scalable & Secure Systems Research (08961), Sandia National Laboratories, Livermore, CA 94550, USA*

(Dated: August 27, 2014)

Precise identification of parameters governing quantum processes is a critical task for quantum information and communication technologies. In this work we consider a setting where system evolution is determined by a parameterized Hamiltonian, and the task is to estimate these parameters from temporal records of a restricted set of system observables (time traces). Based on the notion of system realization from linear systems theory we develop a constructive algorithm that provides estimates of the unknown parameters directly from these time traces. We illustrate the algorithm and its robustness to measurement noise by applying it to a one-dimensional spin chain model with variable couplings.

The promise of quantum technologies for tasks such as computation, communication, and metrology is motivating the construction of devices that are precisely engineered at the nanoscale, and whose quantum dynamics are exceptionally well characterized and controlled [1]. The fragility and sensitivity of typical quantum devices make achieving such objectives extremely challenging, and significant research efforts over the past two decades have focused on addressing these challenges.

Process tomography is the most generally applied technique for characterizing an unknown quantum dynamical process [1, 2]. However, all variants of process tomography are very resource demanding, *e.g.*, in the required number of measurements settings and number of input state preparations. In addition, it is often unsuitable in resource-constrained situations where one may only have measurement access to certain observables or subsystems; *e.g.*, see Fig. 1. Furthermore, process tomography does not utilize often available partial information about the system. One such common scenario is when the structure of a dynamical model can be obtained from underlying physics and what is to be determined are some unknown parameters in the model. This is the quantum version of parameter estimation in classical system sciences, and some previous work has considered variants to quantum tomography for this problem [3].

In this work, we consider a new approach to quantum parameter estimation. Whereas process tomography typically measures a complete basis of system observables at *one* time instant, we ask what can be achieved if a temporal record of a small set of system observables is collected? We refer to such a successive record of observable expectations as an *measurement time trace*, and develop a method that enables information about dynamical parameters to be extracted from such time traces. Our method takes into account *a priori* information and

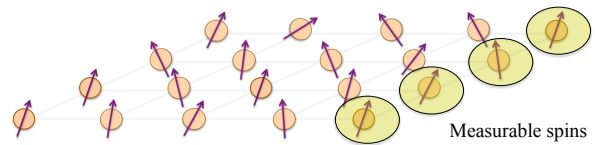


FIG. 1: A spin (or qubit) lattice as an example illustrating the type of system considered in this work. The spins interact with each other through nearest-neighbor or long-range couplings and certain local observables are measurable for a subset of the spins (circled above). The task is to identify the parameters defining the Hamiltonian of the interconnected system from a time trace of expectation values of these observables.

fits naturally into resource constrained situations, and as such we expect that it will be very experimentally relevant and feasible. Additionally, because our scheme utilizes a time trace, it can identify the generator of dynamics (*e.g.*, a Hamiltonian) as opposed to the dynamical map (*e.g.*, a unitary at a fixed time), which is typically what process tomography achieves. This is advantageous since in physically realistic scenarios the generator of dynamics is more compactly specified than the map. This will be discussed in more detail below.

Several authors have considered parameter estimation from various types of time-dependent measurement records [4–14]. Particularly relevant to this work, Cole *et al.* used Fourier analysis to identify a single qubit Hamiltonian from one measurement observable [5], and Devitt *et al.* presented a scheme to identify any two-qubit Hamiltonian from the temporal evolution of the concurrence measure of entanglement [6]. Subsequent work by Burgarth *et al.* [8, 9] and Di Franco *et al.* [10] generalized this approach to estimate the coupling strengths in a many-qubit network from measurements on a small part of the network. Recently, Burgarth *et al.* presented a framework for quantum system identification based on input/output information and formalized the notion of equivalence between system realizations [15].

*Electronic address: mnsarov@sandia.gov

Our approach differs from existing work on quantum system identification in two critical aspects. Firstly, we develop a constructive algorithm for identification of arbitrary Hamiltonian quantum dynamics that takes advantage of available prior knowledge of the system (*e.g.*, network structure or partial knowledge of parameters). The technique can also be employed when such prior information is absent. Secondly, in contrast to most existing system identification schemes, we do not require state tomography of a restricted set of subsystems, but rather develop a technique that produces parameter estimates based only on the collected measurement time traces.

Setup – We consider the task of identifying the Hamiltonian of an unknown quantum dynamical process. Assume that the dimension of the system is finite and known, and that the dynamical process can be prepared at some well-characterized initial states. Further, we assume that the dynamical evolution of the process is unitary (no decoherence). This condition can be relaxed and the approach will be extended to the non-unitary case in a future publication.

A parameterized form of the Hamiltonian governing the quantum dynamical process can be written as,

$$H = \sum_{m=1}^M a_m(\theta) X_m, \quad (1)$$

where θ is a vector consisting of unknown parameters, $a_m \in \mathbb{R}$ are some known functions of θ , and X_m are known Hermitian operators [26]. Assume that the dimension of the quantum process is N , and thus $iH \in \mathfrak{su}(N)$, *i.e.* the Lie algebra consisting of all the $N \times N$ skew-Hermitian matrices. An orthonormal basis of $N^2 - 1$ matrices $\{iX_m\}$ can be chosen for $\mathfrak{su}(N)$, where the Hilbert-Schmidt inner product is defined as $\langle iX_m, iX_n \rangle \equiv \text{tr}(X_m^\dagger X_n)$, and hence $a_m = \text{tr}(HX_m)$. For example, $\frac{i}{2}\sigma_\alpha^1 \otimes \sigma_\beta^2$ form a basis for the two-qubit algebra $\mathfrak{su}(4)$, where $\sigma_\alpha, \sigma_\beta$ can be Pauli matrices $\sigma_x, \sigma_y, \sigma_z$, or the identity matrix I_2 , and superscripts label the qubits [27]. The numbers C_{jkl} such that

$$[iX_j, iX_k] = \sum_{l=1}^{N^2-1} C_{jkl}(iX_l), \quad j, k = 1, \dots, N^2 - 1,$$

are the *structure constants* of the Lie algebra $\mathfrak{su}(N)$ with respect to this basis. Each element X_m is Hermitian and thus can be considered an *observable* for the system. Furthermore, we can consider the a_m as our unknown parameters, because solving for θ from a_m is simply an algebraic problem.

Note that in Eq. (1), typically $M \ll N^2 - 1$ because of physical constraints on system energy, locality, and weight of interactions. For instance, the Hamiltonian for the spin lattice system in Fig. 1 contains only weight-one and weight-two basis elements X_m [28], and furthermore, the weight-two interactions might be restricted to only being between nearest-neighbor spins on the lattice.

By utilizing measurement time traces our identification algorithm can estimate the process at the Hamiltonian level where there are only M unknown parameters. In contrast, process tomography generally does not consider time traces and therefore must estimate the process at the unitary level where there are in general $N^2 - 1$ unknown parameters.

Observable dynamics – The dynamics of the expectation value of an observable X_k , written as $x_k = \langle \psi | X_k | \psi \rangle$, can be derived as

$$\dot{x}_k = \sum_{l=1}^{N^2-1} \left(\sum_{m=1}^M C_{mkl} a_m \right) x_l. \quad (2)$$

Collecting the x_k in a vector $\mathbf{x} \in \mathbb{R}^{N^2-1}$, we obtain a linear equation describing the complete dynamics:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}, \quad x_k(0) = \langle \psi(0) | X_k | \psi(0) \rangle, \quad (3)$$

where the matrix $\mathbf{A} \in \mathbb{R}^{(N^2-1) \times (N^2-1)}$ has elements $\mathbf{A}_{kl} = \sum_{m=1}^M C_{mkl} a_m$. Using the antisymmetries of the structure constants, it can be shown that $\mathbf{A}^\top = -\mathbf{A}$. The vector \mathbf{x} , often called the *coherence vector* [16], is a complete representation of the quantum state. Eq. (12) explicitly describes the quantum dynamics as a linear time invariant (LTI) system and hence it enables application of results from classical linear systems theory.

Typically, some observable expectation values may be easily measured, *e.g.*, local observables of a collection of spins are tracked as function of time, see Fig. 1. Often the measured observables belong to the chosen $\mathfrak{su}(N)$ basis, but if not, each observable O_i can be expanded in this basis as $O_i = \sum_j o_j^{(i)} X_j$. Collect the unique basis elements present in the expansion of all measured observables in the set $\mathcal{M} = \{X_{\nu_1}, X_{\nu_2}, \dots, X_{\nu_p}\}$, where ν is a vector of length p . For example, if $O_1 = o_3^{(1)} X_3 + o_5^{(1)} X_5$ and $O_2 = o_2^{(2)} X_2 + o_3^{(2)} X_3$, with $o_k^{(j)} \in \mathbb{R}$, then $p = 3$ and $\mathcal{M} = \{X_2, X_3, X_5\}$. Generally, $p \ll N^2 - 1$.

In the following we will use time traces of the measured observable expectation values to identify the unknown Hamiltonian parameters. To this end, we first need to derive the dynamical equation governing the time evolution of these observables. Parallel to the study of controllability in classical nonlinear systems theory [17], we give a constructive procedure to obtain the closed dynamics for these observables. For the Hamiltonian in Eq. (1), let $\Delta = \{X_m\}_{m=1}^M$. Define an iterative procedure as

$$G_0 = \mathcal{M}, \text{ and } G_i = [G_{i-1}, \Delta] \cup G_{i-1}, \quad (4)$$

where $[G_{i-1}, \Delta] \equiv \{X_j : \text{tr}(X_j^\dagger [g, h]) \neq 0, \text{ where } g \in G_{i-1}, h \in \Delta\}$ [29]. In geometric control theory, the sequence of G_i are referred to as the *filtration* associated to Δ [17]. Since $\mathfrak{su}(N)$ is finite, this iteration will saturate at a maximal set \bar{G} after finite steps, and we refer to this set as the *accessible set*. Intuitively, the set \bar{G} contains

the elements of the system that couple to the measured observables. Then, writing all the x_k with $X_k \in \bar{G}$ in a vector \mathbf{x}_a of dimension $K \leq N^2 - 1$, the dynamics for this vector is given by

$$\dot{\mathbf{x}}_a = \tilde{\mathbf{A}}\mathbf{x}_a, \quad (5)$$

where $\tilde{\mathbf{A}}$ is a $K \times K$ sub-matrix of \mathbf{A} , *i.e.*, only the elements necessary to describe the evolution of the subset of observable averages collected in \mathbf{x}_a .

Identification algorithm – A necessary condition for the identifiability of a_m is that it be present in the matrix $\tilde{\mathbf{A}}$, because otherwise it would not participate in the dynamical equation (5), and there would be no way to infer its value from examining the observables in \mathcal{M} . In order to estimate these identifiable parameters we utilize the notion of a *system realization* constructed from the measurement time traces. In linear systems theory there are many methods for constructing a realization of a linear dynamical system based on measurement results [18], and in the following we adapt one of these, the *eigenstate realization algorithm* (ERA) [19], for the purposes of Hamiltonian parameter estimation.

The estimation setting we consider is the following. Suppose we have access to the expectation values of the observables in \mathcal{M} at regular time instants $j\Delta t$ for some sampling period Δt [20]. Denote these values as $\{\mathbf{y}(j\Delta t)\}$, and they may have to be collected from averaging measurements on several runs of the experiment under the same initial state. Note that $\mathbf{y}(j\Delta t)$ is the output of the following discretized form of Eq. (5):

$$\mathbf{x}_a(j+1) = \tilde{\mathbf{A}}_d \mathbf{x}_a(j), \quad \mathbf{y}(j) = \mathbf{C}\mathbf{x}_a(j), \quad (6)$$

where for brevity of notation we use $\mathbf{x}_a(j) \equiv \mathbf{x}_a(j\Delta t)$ and $\mathbf{y}(j) \equiv \mathbf{y}(j\Delta t)$, and $\tilde{\mathbf{A}}_d = e^{\tilde{\mathbf{A}}\Delta t}$. The $p \times K$ matrix \mathbf{C} picks up the entries in $\mathbf{x}_a(j)$ that correspond to expectation values of elements of \mathcal{M} . Also assume that the system is prepared at a fixed, known initial state $\mathbf{x}(0)$, and the corresponding initial state for Eq. (6) is $\mathbf{x}_a(0)$. Then these relations can be solved easily to obtain an explicit form for the outputs: $\mathbf{y}(j) = \mathbf{C}\tilde{\mathbf{A}}_d^j \mathbf{x}_a(0)$. Having access to the time trace $\mathbf{y}(j)$, one may try to solve this set of equations directly. However, since $\tilde{\mathbf{A}}_d$ is a transcendental function of a_m , determining the parameters this way is usually infeasible. Instead, we will utilize ERA and formulate a new relationship so that parameter estimation only requires solving polynomial equations.

The first stage of the estimation algorithm is to construct a minimal realization of the system based on input/output information. This is achieved by ERA in three steps, as follows.

Step 1: Collect the measured data into an $rp \times s$

matrix (generalized Hankel matrix) as:

$$\mathbf{H}_{rs}(k) = \begin{bmatrix} \mathbf{y}(k) & \mathbf{y}(k+t_1) & \cdots & \mathbf{y}(k+t_{s-1}) \\ \mathbf{y}(j_1+k) & \mathbf{y}(j_1+k+t_1) & \cdots & \mathbf{y}(j_1+k+t_{s-1}) \\ \vdots & \vdots & & \vdots \\ \mathbf{y}(j_{r-1}+k) & \mathbf{y}(j_{r-1}+k+t_1) & \cdots & \mathbf{y}(j_{r-1}+k+t_{s-1}) \end{bmatrix}$$

with arbitrary integers j_i ($i = 1, \dots, r-1$) and t_l ($l = 1, \dots, s-1$).

Step 2: Find the singular value decomposition (SVD) of $\mathbf{H}_{rs}(0)$ as

$$\mathbf{H}_{rs}(0) = P \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} Q^\top = [P_1 \ P_2] \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} Q_1^\top \\ Q_2^\top \end{bmatrix},$$

where $P \in \mathbb{R}^{rp \times rp}$, $Q \in \mathbb{R}^{s \times s}$ are both orthonormal, and Σ is a diagonal matrix with the non-zero singular values of $\mathbf{H}_{rs}(0)$ determined up to numerical accuracy ϵ , *i.e.*, $\Sigma_{ii} > \epsilon$ for all $i \leq n_\Sigma$ where n_Σ is the dimension of Σ . The matrices P_1 , P_2 , Q_1 , Q_2 are partitions with compatible dimensions.

Step 3: Form a realization of the system (6) as $\hat{\mathbf{A}}_d = \Sigma^{-\frac{1}{2}} P_1^\top \mathbf{H}_{rs}(1) Q_1 \Sigma^{-\frac{1}{2}}$, $\hat{\mathbf{C}} = \mathbf{E}_p^\top P_1 \Sigma^{\frac{1}{2}}$, where $\mathbf{E}_p^\top = [\mathbf{I}_p, 0_p, \dots, 0_p]$. The pair $(\hat{\mathbf{A}}_d, \hat{\mathbf{C}})$ reproduces the input-output relations specified by Eq. (6), that is:

$$\mathbf{y}(j) = \mathbf{C}\tilde{\mathbf{A}}_d^j \mathbf{x}_a(0) = \hat{\mathbf{C}}\hat{\mathbf{A}}_d^j \hat{\mathbf{x}}(0), \quad \text{for all } j \geq 0, \quad (7)$$

provided that $\hat{\mathbf{x}}(0) \equiv \Sigma^{\frac{1}{2}} Q_1^\top e_1$, where e_1 is the first column of \mathbf{I}_s .

This completes the specification of the ERA algorithm. Then let $\hat{\mathbf{A}} = \log \hat{\mathbf{A}}_d / \Delta t$ [20]. This results in a realization of the continuous-time linear system in the form of the triple $(\hat{\mathbf{A}}, \hat{\mathbf{C}}, \hat{\mathbf{x}}(0))$. Now, to estimate the Hamiltonian parameters we use an invariant of different realizations, the *transfer function* [18], to form equations for the unknown parameters. Specifically, the transfer function from an initial state $\mathbf{x}(0)$ to the measurement observables specified by \mathbf{C} can be written as $G(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{x}(0)$, where $s \in \mathbb{C}$ is the Laplace variable. Equating the transfer functions for the original system with unknown parameters and the ERA realization we get:

$$\mathbf{C}(s\mathbf{I} - \tilde{\mathbf{A}})^{-1}\mathbf{x}_a(0) = \hat{\mathbf{C}}(s\mathbf{I} - \hat{\mathbf{A}})^{-1}\hat{\mathbf{x}}(0). \quad (8)$$

The right hand side of Eq. (8) is completely determined by the measured data, and the left hand side can be simplified as the ratio $Q(s)/P(s)$ [18], where

$$P(s) = \det(s\mathbf{I} - \tilde{\mathbf{A}}), Q(s) = \det \left(s \begin{bmatrix} \mathbf{I} & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} \tilde{\mathbf{A}} & \mathbf{x}_a(0) \\ \mathbf{C} & 0 \end{bmatrix} \right). \quad (9)$$

The coefficients of $Q(s)$, $P(s)$ are all polynomials of the Hamiltonian parameters a_m . Equating these coefficients with those in the right hand side of Eq. (8), we obtain

a system of polynomial equations. Solving these multivariate polynomial equations leads to the identification of a_m .

A judicious choice for the initial state is crucial to this identification scheme. For instance, if \mathbf{x}_a is zero or an eigenvector of $\tilde{\mathbf{A}}$, it leads to no sensitivity in the output to any of the unknown parameters. Care must be taken to avoid such degenerate cases. In fact, running the algorithm with multiple initial states leads to more polynomial equations with low order and thus helps to solve these equations more efficiently.

This system identification algorithm can result in multiple estimates of the unknown parameters, all of which satisfy the input/output relations captured by Eq. (8). This is because several system Hamiltonians can generate the same map between an input state and measurement time trace, and hence are equivalent from an input/output perspective [15]. When the algorithm results in multiple parameter estimates and more specification is needed, one has to appeal to prior information, or add resources such as additional input states or observable time traces.

Example – Consider the following Hamiltonian for a one-dimensional chain of n qubits:

$$H = \sum_{k=1}^n \frac{\omega_k}{2} \sigma_z^k + \sum_{k=1}^{n-1} \delta_k (\sigma_+^k \sigma_-^{k+1} + \sigma_-^k \sigma_+^{k+1}).$$

This Hamiltonian is often used as a model for a spin “wire” that enables quantum state transfer [21]. Suppose that only one end of the spin chain is observable, and choose $\langle \sigma_x^1 \rangle$ as the observable that is tracked. Choosing the generalized Pauli operators as our basis and calculating the filtration per Eq. (4) yields the accessible set as $\tilde{G} = \{2^{-n/2} \sigma_x^1, 2^{-n/2} \sigma_y^1\} \cup \{2^{-n/2} \sigma_z^1 \dots \sigma_z^{k-1} \sigma_x^k, 2^{-n/2} \sigma_z^1 \dots \sigma_z^{k-1} \sigma_y^k\}_{k=2}^n$. The system matrix $\tilde{\mathbf{A}}$ is $2n \times 2n$ and has the following simple structure

$$\tilde{\mathbf{A}} = \begin{bmatrix} 0 & \omega_1 & 0 & -\delta_1 & & & & \\ -\omega_1 & 0 & \delta_1 & 0 & 0 & & & \\ 0 & -\delta_1 & 0 & \omega_2 & 0 & \ddots & & \\ \delta_1 & 0 & -\omega_2 & 0 & \ddots & \ddots & 0 & \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & -\delta_{n-1} \\ & \ddots & \ddots & \ddots & 0 & \delta_{n-1} & 0 & \\ & & 0 & 0 & -\delta_{n-1} & 0 & \omega_n & \\ & & \delta_{n-1} & 0 & -\omega_n & 0 & & \end{bmatrix}$$

with $\mathbf{x}_a = [\bar{x}_1, \bar{y}_1, \dots, \bar{x}_n, \bar{y}_n]$, where $\bar{x}_1 \equiv \langle \sigma_x^1 \rangle$, $\bar{y}_1 \equiv \langle \sigma_y^1 \rangle$ and $\bar{x}_k \equiv \langle \sigma_z^1 \dots \sigma_z^{k-1} \sigma_x^k \rangle$, $\bar{y}_k \equiv \langle \sigma_z^1 \dots \sigma_z^{k-1} \sigma_y^k \rangle$ for $k \geq 2$. In this basis $\mathbf{C} = [1, 0, \dots, 0]$. All parameters in the Hamiltonian appear in $\tilde{\mathbf{A}}$, and therefore the necessary condition for identifying all parameters is satisfied for an estimation strategy that uses only time traces of $\langle \sigma_x^1 \rangle$.

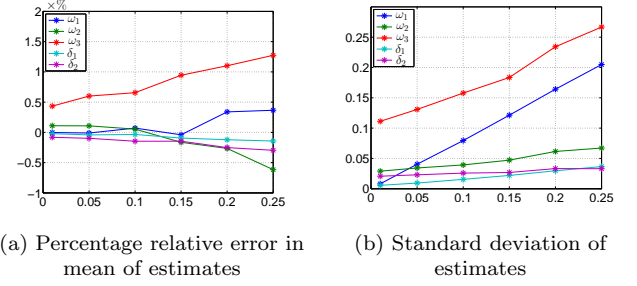


FIG. 2: Assessing the robustness of parameter estimation algorithm. The x -axis in both figures is the standard deviation of the measurement noise, σ .

Choosing an initial state $\frac{|0\rangle + i|1\rangle}{\sqrt{2}}|0 \dots 0\rangle$ (with corresponding coherence vector $[0, 1, 0, \dots, 0]^T$), and running ERA results in a realization $(\hat{\mathbf{A}}, \hat{\mathbf{C}}, \hat{\mathbf{x}}(0))$. The transfer function is given by

$$\mathbf{C}(s\mathbf{I} - \tilde{\mathbf{A}})^{-1} \mathbf{x}_a(0) = \frac{q_{2n-2}s^{2n-2} + \dots + q_2s^2 + q_0}{s^{2n} + p_{2n-2}s^{2n-2} + \dots + p_2s^2 + p_0},$$

where the detailed expressions of the coefficients p_i and q_i as polynomials of ω_k and δ_k can be calculated via Eq. (11). These equations can be solved by mature numerical toolboxes such as PHCpack [22] to obtain the unknown parameters ω_k and δ_k . In the Supplementary Material we simulate time traces for this model with $n = 3$ and solve these polynomial equations to explicitly demonstrate the parameter estimation algorithm [20]. In the absence of measurement noise, the parameters can be perfectly identified up to sign of δ_k . The sign ambiguity is because the coupling strengths only occur to even order in the polynomial equations when the local observable being measured is $\langle \sigma_x^1 \rangle$. Additional measurements or prior information are required to determine the sign.

Experimental measurements of observable expectation values will inevitably be noisy, and therefore we also assess the performance of our estimation algorithm in the presence of measurement noise. Consider the case where the measurements in the 3-qubit example specified in the Supplementary Material are corrupted by additive Gaussian noise, *i.e.*, $\mathbf{y}(j) = \langle \sigma_x^1 \rangle(j) + \eta(j)$, with $\eta(j) \sim \mathcal{N}(0, \sigma)$. The observable $\langle \sigma_x^1 \rangle(j)$ lies in the range $[-1, 1]$, and we consider noise with σ values 0.01, 0.05, 0.10, 0.15, 0.20, and 0.25. For each σ , we generate 4000 Gaussian noise trajectories and estimate the five parameters, $\theta = (\omega_1, \omega_2, \omega_3, \delta_1, \delta_2)$, from each noisy measurement trace. Fig. 2 shows summary statistics that demonstrate the accuracy and robustness of the estimation procedure. The relative error in the mean of the parameter estimates, $\frac{\bar{\theta}_i - \theta_i}{\theta_i} \times 100\%$ [30], remains small, whereas the standard deviation of the estimates scales approximately linearly with σ . Further characterization of the robustness of the procedure to measurement noise is presented

in the Supplementary Material. We note that the robustness of our method is a function of the realization algorithm (ERA) and realization invariant used to construct the polynomial equations. In fact, we experimented with another invariant, the *Markov parameters* of a system, and discovered that it is not as robust to noise as the transfer function approach presented here.

Conclusion – We have developed a robust algorithm to identify the unknown parameters of a quantum Hamiltonian from the time traces of a set of system observables, which naturally takes into account prior information and restrictions on measurement access. A direction for future work is the generalization of this algorithm to parameter estimation for open quantum systems governed by Lindblad evolution [23], in which case the evolution of the coherence vector is described by an affine time-invariant system of equations [16].

Acknowledgments – MS thanks Akshat Kumar for information on techniques for solving multivariate polynomial systems. This work was supported by the Laboratory Directed Research and Development program at Sandia National Laboratories. Sandia is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the United States Department of Energy’s National Nuclear Security Administration under contract DE-AC04-94AL85000. JZ acknowledges financial support from NSFC under Grant No. 61174086, and State Key Laboratory of Precision Spectroscopy, ECNU, China. The authors are grateful for the hospitality of KITP at UCSB, where this work was initiated. This research was supported in part by the National Science Foundation under Grant No. NSF PHY11-25915.

[1] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum info.* (Cambridge University Press, 2001).
[2] M. Mohseni, A. Rezakhani, and D. Lidar, Phys. Rev. A **77**, 032322 (2008).
[3] M. P. da Silva, O. Landon-Cardinal, and D. Poulin, Phys. Rev. Lett. **107**, 210404 (2011).
[4] N. Boulant, T. Havel, M. Pravia, and D. Cory, Phys. Rev. A **67**, 042322 (2003).
[5] J. H. Cole, S. G. Schirmer, A. D. Greentree, C. J. Wellard, D. K. L. Oi, and L. C. L. Hollenberg, Phys. Rev. A **71**, 062312 (2005).
[6] S. J. Devitt, J. H. Cole, and L. C. L. Hollenberg, Phys. Rev. A **73**, 052317 (2006).
[7] K. C. Young, M. Sarovar, R. Kosut, and K. B. Whaley, Phys. Rev. A **79**, 062301 (2009).
[8] D. Burgarth, K. Maruyama, and F. Nori, Phys. Rev. A **79**, 020305 (2009).
[9] D. Burgarth and K. Maruyama, New J. Phys. **11**, 103019 (2009).
[10] C. Di Franco, M. Paternostro, and M. S. Kim, Phys. Rev. Lett. **102**, 187203 (2009).
[11] C. E. Granade, C. Ferrie, N. Wiebe, and D. G. Cory,

New J. Phys. **14**, 103013 (2012).
[12] Y. Kato and N. Yamamoto, in *Proc. 52nd IEEE CDC* (2013), arXiv.org: 1305.0418.
[13] J. M. Dominy, L. C. Venuti, A. Shabani, and D. A. Lidar, arXiv.org (2013), 1312.0973v1.
[14] V. Jagadish and A. Shaji, arXiv.org (2014), 1401.1058v1.
[15] D. Burgarth and K. Yuasa, Phys. Rev. Lett. **108**, 080502 (2012).
[16] K. Lendi, J. Phys. A: Math. Theor. **20**, 15 (1987).
[17] S. Sastry, *Nonlinear systems* (Springer, 1999).
[18] F. M. Callier and C. A. Desoer, *Linear System Theory* (Springer New York, New York, NY, 1991).
[19] J. N. Juang and R. S. Pappa, Journal of Guidance, Control, and Dynamics **8**, 620 (1985).
[20] See Supplementary Material for discussion of appropriate choice of Δt and an application of the system identification algorithm..
[21] S. Bose, Contemporary Physics **48**, 13 (2007).
[22] Y. Guan and J. Verschelde, *PHClab: A MATLAB/Octave interface to PHCpack* (Springer-Verlag, 2008), vol. 148 of *IMA: Software for Alge. Geom.*
[23] H.-P. Breuer and F. Petruccione, *The theory of open quantum systems* (Oxford University Press, 2002).
[24] G. F. Franklin, J. D. Powell, and M. L. Workman, *Digital Control of Dynamic Systems* (Addison-Wesley, 1997).
[25] N. J. Higham, *Functions of matrices: theory and computation* (SIAM, 2008).
[26] We set $\hbar = 1$ and therefore the a_m have units $1/s$.
[27] In the following we will omit the tensor product when writing multi-qubit Pauli operators for brevity.
[28] The *weight* of a multi-qubit Pauli operator is the number of non-identity terms in the tensor product.
[29] We do not need to keep track of multiplicative constants, only the operators generated by these commutators.
[30] \bar{X} is the empirical mean of the random variable X .

SUPPLEMENTARY INFORMATION FOR “QUANTUM HAMILTONIAN IDENTIFICATION FROM MEASUREMENT TIME TRACES”

Choosing the sampling period

The starting point for our system identification algorithm is a time trace representing sampled outputs of the system. Forming the realization $\hat{\mathbf{A}}$ is equivalent to reconstructing the continuous time system, and therefore we expect that a judicious choice of sampling period Δt in the original time trace is important to obtain accurate results from the algorithm. In this section we outline the requirements for Δt .

From $\hat{\mathbf{A}}^\top = -\hat{\mathbf{A}}$, we know that the eigenvalues of $\hat{\mathbf{A}}$ are all pure imaginary numbers. Therefore the observable dynamics determined by $\dot{\mathbf{x}}_a = \hat{\mathbf{A}}\mathbf{x}_a$ is a summation of sinusoidal functions, whose frequencies are given by the eigenvalues of $\hat{\mathbf{A}}$. To perfectly recover the continuous time dynamics, we will require the sampling time Δt to satisfy the Nyquist Sampling Theorem [24], which states that the sampling frequency needs to be at least twice the highest frequency in the observable dynamics. The highest angular frequency is given by $\max |\sigma(\hat{\mathbf{A}})|$, where $\sigma(\hat{\mathbf{A}})$ denotes the spectrum of $\hat{\mathbf{A}}$. This in turn yields the corresponding highest frequency as $\max |\sigma(\hat{\mathbf{A}})|/2\pi$. Hence, Nyquist Sampling Theorem imposes a requisite condition on the sampling frequency f_{sampling} as

$$f_{\text{sampling}} > 2 \frac{\max |\sigma(\hat{\mathbf{A}})|}{2\pi}, \quad (10)$$

which leads to

$$\Delta t = \frac{1}{f_{\text{sampling}}} < \frac{\pi}{\max |\sigma(\hat{\mathbf{A}})|}. \quad (11)$$

Eq. (11) is a condition on how to choose the sampling period Δt , and the right hand side is a time scale describing the system. Of course, in a Hamiltonian parameter estimation problem, we usually do not know the eigenvalues of the matrix $\hat{\mathbf{A}}$. Hence we will need to guess a suitable sampling time and then refine it with an adaptive method if necessary.

Choosing a sampling period satisfying Eq. (11) becomes particular important when taking the matrix logarithm of $\hat{\mathbf{A}}_d$. A sampling period less than required by the Nyquist Sampling Theorem implies that this logarithm is defined uniquely. To see this, note that Eq. (11) implies

$$\max |\sigma(\hat{\mathbf{A}}\Delta t)| < \pi. \quad (12)$$

Since $(\hat{\mathbf{A}}, \hat{\mathbf{C}}, \hat{\mathbf{x}}(0))$ from ERA is a minimal realization of the original system represented by $(\tilde{\mathbf{A}}, \mathbf{C}, \mathbf{x}_a(0))$, the eigenvalues of $\hat{\mathbf{A}}$ must also be the eigenvalues of $\tilde{\mathbf{A}}$. Therefore, we obtain

$$\max |\sigma(\tilde{\mathbf{A}}\Delta t)| < \pi. \quad (13)$$

Now let us quote the following Theorem from Page 20 in Ref. [25], which introduces the notion of principal logarithm:

Theorem 1.31 Let $A \in \mathbb{C}^{n \times n}$ have no eigenvalues on \mathbb{R}^- . There is a unique logarithm X of A all of whose eigenvalues lie in the strip $\{z : -\pi < \text{Im}(z) < \pi\}$. We refer to X as the principal logarithm of A and write $X = \log(A)$. If A is real then its principal logarithm is real.

Here \mathbb{R}^- denotes the negative real axis. Since the eigenvalues of $\hat{\mathbf{A}}$ are purely imaginary, bounded as Eq. (13), and $\hat{\mathbf{A}}_d = e^{\hat{\mathbf{A}}\Delta t}$, we know that $\hat{\mathbf{A}}_d$ has no eigenvalues on \mathbb{R}^- . Therefore, the Theorem above applies and there is a unique principal logarithm. Furthermore, note that the conclusion that the eigenvalues of the principal logarithm lie in the strip $\{z : -\pi < \text{Im}(z) < \pi\}$ is consistent with the properties of $\hat{\mathbf{A}}$ provided that the sampling time is sufficiently small so as to satisfy Eq. (13).

Therefore we see that the accuracy of the algorithm relies on the sampling time of the measurement time trace being sufficiently small.

Example: three qubit XX spin chain

In this section we explicitly demonstrate our system identification algorithm for the spin chain example in the main text, with $n = 3$ qubits.

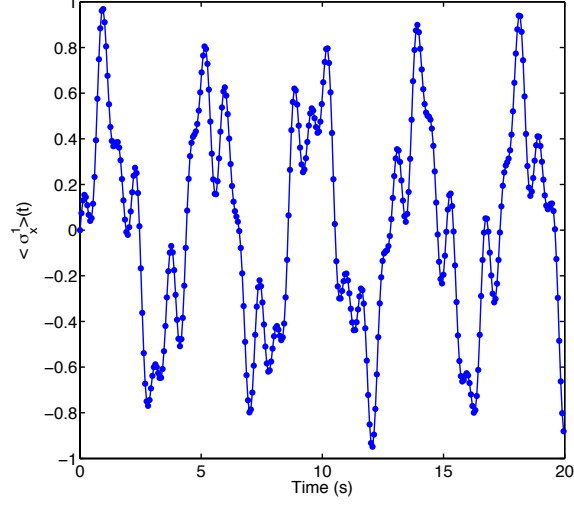


FIG. 3: Measurement time trace for observable $\langle \sigma_x^1 \rangle$ for the XX spin chain example with $n = 3$ qubits. The dots show a sampled measurement trace for initial state $\frac{|0\rangle + i|1\rangle}{\sqrt{2}}|00\rangle$.

Consider the following Hamiltonian:

$$H = \sum_{k=1}^3 \frac{\omega_k}{2} \sigma_z^k + \sum_{k=1}^2 \delta_k (\sigma_+^k \sigma_-^{k+1} + \sigma_-^k \sigma_+^{k+1}), \quad (14)$$

with nominal true parameter values $\omega_1 = 1.3$, $\omega_2 = 2.4$, $\omega_3 = 1.7$, $\delta_1 = 4.3$, $\delta_2 = 5.2$ (all parameters have units 1/sec).

Choose the initial state as $\frac{|0\rangle + i|1\rangle}{\sqrt{2}}|0 \cdots 0\rangle$, and the corresponding coherence vectors $\mathbf{x}_a(0)$ is $[0, 1, 0, \cdots, 0]^T$. Let us assume that we measure the observable $\langle \sigma_x^1 \rangle$ as in the main text. The Laplace transform of the output $\mathbf{y}(t)$ can be written as

$$\mathbf{Y}(s) = \mathbf{C}(s\mathbf{I} - \tilde{\mathbf{A}})^{-1} \mathbf{x}_a(0) = \frac{q_4 s^4 + q_2 s^2 + q_0}{s^6 + p_4 s^4 + p_2 s^2 + p_0}, \quad (15)$$

where

$$\begin{aligned} p_4 &= 2\delta_1^2 + 2\delta_2^2 + \omega_1^2 + \omega_2^2 + \omega_3^2, \\ p_2 &= \delta_1^4 + 2\delta_1^2\delta_2^2 - 2\delta_1^2\omega_1\omega_2 + 2\delta_1^2\omega_3^2 + \delta_2^4 - 2\delta_2^2\omega_2\omega_3 + 2\delta_2^2\omega_1^2 + \omega_1^2\omega_2^2 + \omega_2^2\omega_3^2 + \omega_1^2\omega_3^2, \\ p_0 &= \delta_1^4\omega_3^2 + 2\delta_1^2\delta_2^2\omega_1\omega_3 - 2\delta_1^2\omega_1\omega_2\omega_3^2 + \delta_2^4\omega_1^2 - 2\delta_2^2\omega_1^2\omega_2\omega_3 + \omega_1^2\omega_2^2\omega_3^2, \end{aligned} \quad (16)$$

and

$$\begin{aligned} q_4 &= \omega_1, \\ q_2 &= \omega_1\omega_2^2 - \delta_1^2\omega_2 + 2\delta_2^2\omega_1 + \omega_1\omega_3^2, \\ q_0 &= -\delta_1^2\omega_2\omega_3^2 + \omega_1\omega_2^2\omega_3^2 - 2\delta_2^2\omega_1\omega_2\omega_3 + \delta_2^4\omega_1 + \delta_1^2\delta_2^2\omega_3. \end{aligned} \quad (17)$$

Fig. 3 shows measurement time traces for the initial state $\frac{|0\rangle + i|1\rangle}{\sqrt{2}}|00\rangle$ when simulated for $T = 20s$ with $\Delta t = 0.0598s$. Using this data we construct the Hankel matrix $\mathbf{H}_{rs}(0)$ with $r = 167$, $s = 167$ and all $j_i = 1$ and $t_l = 1$. Then, performing the remaining ERA steps we obtain a realization $(\hat{\mathbf{A}}_d, \hat{\mathbf{C}}, \hat{\mathbf{x}}(0))$. Further taking the logarithm results in a realization $(\hat{\mathbf{A}}, \hat{\mathbf{C}}, \hat{\mathbf{x}}(0))$ of the continuous system. This realization has the same dimension as the original $\tilde{\mathbf{A}}_d$, i.e., $n_\Sigma = 6$.

For the parameter estimation stage of the algorithm we need to pick the five lowest order polynomial equations

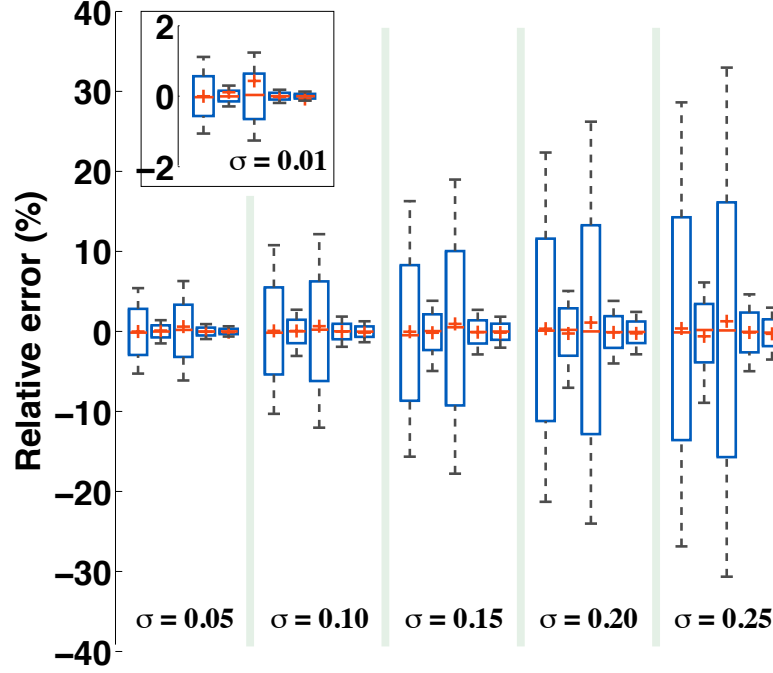


FIG. 4: Box plots for estimates formed from noisy measurement records. Each group of box plots is for estimation using measurement records with noise of standard deviation indicated by σ . The inset shows box plots for $\sigma = 0.01$ separately since the range of relative error in parameter estimates in this case is much smaller than for the other cases. The five box plots in each group are for estimates of parameters (from left to right): $\omega_1, \omega_2, \omega_3, \delta_1, \delta_2$.

from Eqs. (16) and (17) (since there are five unknown parameters in this system):

$$\begin{aligned}
 \omega_1 &= 1.3 \\
 2\delta_1^2 + 2\delta_2^2 + \omega_1^2 + \omega_2^2 + \omega_3^2 &= 101.4 \\
 \omega_1\omega_2^2 - \delta_1^2\omega_2 + 2\delta_2^2\omega_1 + \omega_1\omega_3^2 &= 37.173 \\
 \delta_1^4 + 2\delta_1^2\delta_2^2 - 2\delta_1^2\omega_1\omega_2 + 2\delta_1^2\omega_3^2 + \delta_2^4 - 2\delta_2^2\omega_2\omega_3 + 2\delta_2^2\omega_1^2 + \omega_1^2\omega_2^2 + \omega_2^2\omega_3^2 + \omega_1^2\omega_3^2 &= 1966.4892 \\
 -\delta_1^2\omega_2\omega_3^2 + \omega_1\omega_2^2\omega_3^2 - 2\delta_2^2\omega_1\omega_2\omega_3 + \delta_2^4\omega_1 + \delta_1^2\delta_2^2\omega_3 &= 1407.01176
 \end{aligned}$$

These equations can be solved by mature numerical toolboxes such as **Singular**, **Macaulay 2**, **SOSTools**, and **PHCpack**. In particular, we applied **PHCpack** [22] to obtain the following estimates for the parameters:

$$\hat{\omega}_1 = 1.3, \hat{\omega}_2 = 2.4, \hat{\omega}_3 = 1.7, \hat{\delta}_1 = \pm 4.3, \hat{\delta}_2 = \pm 5.2.$$

The estimates exactly match the true parameters, except for the indeterminate sign for the coupling parameters. As discussed in the main text, this uncertainty in the sign is a result of the equivalence of systems under some input/output maps, and cannot be resolved unless additional measurements and/or initial states are introduced.

Robustness to noise

As shown in the main text, the Hamiltonian parameter estimation algorithm we have developed is robust to measurement noise. To demonstrate this, we perturbed the measurement of observable $\langle \sigma_x^1 \rangle$ with additive Gaussian noise trajectories; *i.e.*, $\mathbf{y}(j) = \langle \sigma_x^1 \rangle(j) + \eta(j)$, with $\eta(j) \sim \mathcal{N}(0, \sigma)$. We consider noise with σ values 0.01, 0.05, 0.10, 0.15, 0.20, and 0.25. For each σ , we generate 4000 Gaussian noise trajectories and estimate the five parameters, $\theta = (\omega_1, \omega_2, \omega_3, \delta_1, \delta_2)$, from each noisy measurement trace.

Fig. 2 in the main text shows the mean and standard deviation of the 4000 estimates for each parameter. To further characterize the variation of the estimates in that figure, we also shows box plots for the relative error in estimates, in

Fig. 4. The red + in each box plot indicates the mean of the estimates. The red line in each box indicates the median while the bottom and top of each box indicate the 25th and 75th percentile of the data, respectively. The end points of the whiskers represent the 9th and 91st percentiles. Interestingly, some parameter estimates are more sensitive to noise than others. In this example, $\hat{\omega}_1$ and $\hat{\omega}_3$ are the most sensitive.

Note on n_Σ

The rank of Hankel matrix, n_Σ , is the size of the reconstructed realization $\hat{\mathbf{A}}_d$ (or $\hat{\mathbf{A}}$), and is an informative parameter. If $n_\Sigma < K$, this means that the original dynamical systems lacks complete controllability or observability. An obvious way in which this can happen is if, for example, some coupling parameters for a network of qubits are actually zero and thus part of the network is decoupled from the portion being measured. That part of the system is then irrelevant for the dynamics captured in the Hankel matrix and is non-identifiable from the measured observables.